DFKI Approach to Dialogue Management

Norbert Reithinger, Elsa Pecourt, Markus Löckelt
{bert,loeckelt,pecourt}@dfki.de
What we develop at DFKI

• Research goal:
  – Develop multimodal interactivity interfaces

• Hypotheses:
  – Multimodal interaction facilitates interaction
  – Analysis, dialog management and presentation are tightly connected

• Major task in our group:
  – Design and implement (parts of) the dialog backbone of such systems

• Design approach (our mantras)
  – Use general solutions for general problems
  – No processing and presentation without representation
Dialogue Management

Complex interplay of

- **Action Planner** – the mind of the machine:
  - knows all other acting parties and how to handle them
  - Interaction Types:
    - Request/Response, e.g., DB query, user interaction
    - Active processes, e.g., route guidance

- **Discourse Model(ler)** – the memory of the machine

- Analysis/Presentation modules
- Functionalities

Current emphasis:

- Flexible and adaptable to many functionalities
- Switch between applications during dialogs
- Reusable (comes with a price)
System 1: SmartKom

Multimodal Dialogue Backbone

- **Application Layer**
  - **Mobile:** Car and Pedestrian Navigation, Tourist Information
  - **Public:** Cinema, Phone, Fax, Mail, Biometrics
  - **Home:** Consumer Electronics, EPG

**SmartKom-Mobile:** Mobile Travel Companion

**SmartKom-Home:** Infotainment Companion

**SmartKom-Public:** Communication Companion
# SmartKom

3 scenarios  
14 applications  
52 functionalities

3500 words at start  
3 PCs (2xLinux, 1x Windows)

(... and 12 partner institutions collaborating)

(Source: Reithinger et al: SmartKom - Adaptive and Flexible Multimodal Access to Multiple Applications. In ICMI '03)
Module Overview of SmartKom

Common Ontology/M3L Schema

Expectations

Modules related to discourse management
Process Model

General Process
- Abstract Duplication Process
- Abstract Imitation Process
- Abstract Repetition Process
- Abstract Replacement Process
- Abstract Reset Process

Mental Process
- Cognitive Process
  - Information Search Process
  - Planning Process

Emotion Process
- Emotion Active Process
- Emotion Directed Process
- Emotion Experience Subject Process
- Emotion Experience Object Process

Perceptual Process
- Hear Perceptual Process
- Watch Perceptual Process

Physical Process

Social Process

Controlling Process
- Controlling Device Process
- Controlling Entertainment Device
- Controlling Presentation Device

Communicative Process

Instructive Process

Motion Process

Presentation Process

Static Spatial Process

Transaction Process
System 2: MIAMM – Speech and Haptics

- Application: Search in MP3 Database (DE, FR, EN)
- Haptic buttons enable new interaction
- Simulated with Phantom devices
- Other visualisation metaphors
Modules of MIAMM

- Speech Recognition
- Speech Analysis
- Visual / haptic / tactile Interaction
- MP3-Player
- Speech-Prompts
- Common Rep. (MMIL)
- Dialog Manager
- Multimodal-Fusion
- Dialog-Memory
- Action-Planner
- Domain-Model Music Database
Adaptation Needed

- **Representation/Ontology:**
  - Used by almost all modules
    - Provided by partners (SmartKom and MIAMM)
    - Basis of reasoning in the Discourse Modeller (SmartKom)

- Interpretation rules in **Speech Interpretation** (SPIN) to create semantic representations
- Plan operators in the **Action Planner**
- **Interface** to the backend application/application logic
  - E.g., we cannot work very well with the prompts/text in the backend

- Functionality that could be skipped:
  - Presentation plans for the **Presentation Planner**
  - Generator rules for the **Text Generator**

- Adaptation time per application in SmartKom: approx. 1-2 person months (2-3 weeks, 2-3 persons)
Approach in the Action Planner: Communication Channels

• Unified treatment of all communication channels (user modalities, internal and external functionalities)

• Input and output channels defined for every module communicating with AP.
  – In MIAMM: MMF, VisHapTac, MiaDoMo, MP3 Player
  – In SmartKom: Intention Recognition, Discourse Modeller, Dynamic Help, Presentation Planning, Function Modelling

• Processing Approach: Communicative Games
Communicative Games (1)

• **Communicative Games** (CGs) for the modelling of dialogues (plan operators)
  – A CG consists of a sequence of communication acts (dialogue moves) e.g. *request*, *response*, *inform*, *greeting*
  – Can be embedded in other CG
  – synchronous and asynchronous communication possible

• The Action Planner
  – Creates a plan made of CGs, to achieve the goals of the user,
  – triggers and controls the execution of the actions
Communicative Games (MIAMM)

Interaction

Query

User

Presentation

request

response

„I want something from the 90’s“

„Here are the results“
Communicative Games

Interaction

User

Extern

Presentation

MiaDoMo (Database)

Query

Reply

request

response

query

results
Communicative Games

Interaction

User

External

Presentation

Query

Database

Clarification dialogue (System initiative)

request

response

Question

Answer

Query

Reply

“What kind of music are you looking for”
Plan operators in MIAMM

I want something from the 90's

Here are the results
Plan operators in MIAMM (2)
Planing Approach: Backward Chaining

Current State
- provides(?)
- needs(?)
- provides(F,G)
- needs(H)

Successor States
- provides(C)
- needs()
- provides(A)
- needs(C)
- provides(B)
- needs(D)
- provides(D)
- needs(E)
- provides(Y)
- needs(D)
  - condition(D=x)
- provides(X)
  - needs(A,B)
- provides(X)
  - needs()
- provides(X)
  - needs(D)
- provides(A)
  - needs(C)
- provides(B)
  - needs(D)
- provides(C)
  - needs()

Goal
- provides()
  - needs(X,Y)
Pros and Cons of Our Approach

Pros:
• Generic Methods for
  – Analysis/Generation
  – Reference processing
  – Action Planning
• Integration of multimodal I/O
• Neutral to functionalities
• Straightforward method of adaptation

Cons:
• Needs deep(-er) modeling of the domain
• Lots of modules need to go along for a new application
• Difficult to integrate functionalities like the WS-backend which contains e.g. answer texts
• Separation of Dialog History and Action Planner
What Would We Like to Cover

• Error handling in the planner:
  – MIAMM: straightforward solution as plan operators
  – SmartKom: separate Dynamic Help module (due to reasons of project structure)
• Less applications, more generic dialog phenomena
  – More explicit feedback about the systems intentions and the ongoing dialog to the user
  – Dialog processing better adapted to the user
  – Knowledge about argumentations/discussions/negotiations (current bias on access to functionalities)

• Questions:
  – Which abstractions are really needed?
  – Which abstractions are useful?
Thank you very much for your attention!

- Multimodal projects with participation of DFKI
  - SmartKom (BMBF): [http://www.smartkom.org](http://www.smartkom.org)